

Cerebrovascular Diseases Dashboard and Update of the Predictive Analysis: A Proposal for e-SUS Brazil in the Primary Care

Grace TM DAL SASSO^{a,1}, Fades F HAMMES^b and Lucas POSTAL^b

^a*Universidade Federal de Santa Catarina, Brasil*

^b*Laboratório Bridge-UFSC, Brasil*

Keywords. Logistic regression, predictive analysis, cerebrovascular diseases, health informatics, public health

1. Introduction

Information and communication technology and the eHealth strategy propelled data capture and the need to transform them into information with agility and security so that it is possible to generate subsidies for decision making [1].

One of the goals of the Sustainable Development of the United Nations is to reduce the rate of premature mortality from Noncommunicable Chronic Diseases by one-third by 2030. Specifically, cerebrovascular diseases are the second leading cause of death worldwide, behind ischemic heart diseases only. Intelligent tools capable of predicting the probability of a citizen developing cerebrovascular diseases are essential tools for reducing these diseases. Predicting cases that will require qualified care with acceptable accuracy enables acting to prevent clinical problems, reducing costs with expensive treatments, and directing public health investments by making strategies to intervene in advance [2,3].

This study aimed to develop a preliminary predictive model capable of predicting the probability of a citizen developing cerebrovascular diseases in Brazil. The input data for this model was collected through the citizens' electronic health record system and transmitted from one installation to another through the layout of data and interface.

2. Approach

An innovative quantitative study of technological and analytical production used a regression optimized by an algorithm called Stochastic Gradient Descent algorithm (SGD) as a predictive analysis model proposed to estimate the probability of citizens developing cerebrovascular diseases in Primary Health Care.

¹ Corresponding Author, Grace TM Dal Sasso, Av. Governador Ivo Silveira 177, ap502 Estreito, Florianópolis, SC, Brazil. CEP: 88075005.

The artefact created was an API for automation in importing data from the eSUS system base in primary health care for the Cerebrovascular Disease Dashboard (CDD). In addition, we added to the database new structures that allowed the inclusion of information on diets with high concentrations of sodium and sugar and the information on the educational level of citizens.

We used the principles and foundations of Design Science Research Methodology (DSRM) to guide the development of this study. The DSRM is organized into six steps. After this, the data used for the predictive model were randomly generated fictitiously and generated from the registry variables in primary health care. And, the prediction used empirical data from eSUS in primary care through the eSUS Primary Care Data and Interface Layout, and the results were analyzed using descriptive statistics.

3. Body

We carried out the algorithm's training with the registers whose variable cerebrovascular disease had a value of one. In this way, the algorithm identified the factors that contributed to the development of cerebrovascular diseases. Also included in the sample were older citizens, healthy eating habits, and higher education without the condition. This set aimed to inform the algorithm which characteristics led to a low chance of having the disease. The sample used in training totaled 13,561 citizens. Finally, the remaining data set were analyzed and estimated the probability that a citizen will develop this clinical situation.

The inclusion of the variables diet with a high concentration of sodium, diet with a high concentration of sugar, and years of study to the predictive analysis algorithm resulted in an increase in the average probability of a citizen developing the cerebrovascular disease by 45.32%, from 47.59% to 69.16%. However, in 5,452 citizens, the chance decreased, and in 798, the result was the same in both analyzes. These numbers represent 3.20% and 0.47% of the population used in this study, respectively. Higher education and Cerebrovascular disease variables differed from the general average (69.16%), with an average probability of 57.80% and 81.52%, respectively. The minimum possibilities fluctuated considerably. Again, higher education and Cerebrovascular disease variables showed extreme percentages, 19.00% and 38.00%, respectively. The variables pregnant and primary education had a minimum probability of 36.00%, the second-highest score in the category.

The variable higher education had the lowest probabilities, 19.00% for minimum, 88.00% for maximum, and 57.80% on average. On the other hand, the primary education variable had the highest index, 95.00%. Diets with high concentrations of sodium and sugar showed minimum rates of 31.00% and 28.00%, respectively. Regarding the highest indexes, both variables reached the maximum percentage of the prediction, 95.00% of probability. The averages of variables mentioned above, the general standard was 73.00% for a diet with a high concentration of sugar and 72.15% for a diet with a high sodium concentration.

4. Conclusions

The development of this study increased the algorithm confidence index used in predictive analytics, resulting in the likelihood of citizens develop cerebrovascular

diseases and allowed the integration with the electronic health system in primary care. It highlights the importance of analyzing the variables education, a diet with a high concentration of sugar, and a diet with a high sodium concentration in chronic non-communicable diseases. It should be noted that this model developed must be used as an instrument to support citizen care. It should not be used as the sole source of alert for the risk of cerebrovascular diseases. Under no circumstances, the model covers the set of factors and clinical assessments carried out by health professionals and, therefore, further studies are needed.

References

- [1] Bartley A. (2017) Predictive analytics in healthcare: A data-driven approach to transforming care delivery. White Paper. Intel Corporation. 2017. 01-08. <https://01.org/sites/default/files/downloads/model-predicting-rapid-response-events/intelpredictiveanalyticswhitepaperfinal.pdf>
- [2] Parmar P et al. (2015) The Stroke Riskometer™ app: Validation of a data collection tool and stroke risk predictor. Stroke Riskometer™ Collaboration. International Journal of Stroke, 10(2), 231–244. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4335600/>
- [3] Yang L, Wu H, Jin X et al. (2020). Study of cardiovascular disease prediction model based on random forest in eastern China. Sci Rep 10 (5245). <https://doi.org/10.1038/s41598-020-62133-5>